

1656

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/542,520

DATE: 05/31/2001

TIME: 11:51:53

Input Set : A:\7969.076.999.txt

Output Set: C:\CRF3\05312001\I542520.raw

p.s

ENTERED

3 <110> APPLICANT: Jackson, W. James  
4 Pace, John L.  
6 <120> TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF  
8 <130> FILE REFERENCE: 7969-076-999  
10 <140> CURRENT APPLICATION NUMBER: 09/542,520  
11 <141> CURRENT FILING DATE: 2000-04-03  
13 <150> PRIOR APPLICATION NUMBER: PCT/US98/20737  
14 <151> PRIOR FILING DATE: 1998-10-01  
16 <160> NUMBER OF SEQ ID NOS: 41  
18 <170> SOFTWARE: PatentIn Ver. 2.0  
20 <210> SEQ ID NO: 1  
21 <211> LENGTH: 4435  
22 <212> TYPE: DNA  
23 <213> ORGANISM: Artificial Sequence  
25 <220> FEATURE:  
26 <223> OTHER INFORMATION: Description of Artificial Sequence:  
27 recombinant expression vector  
29 <400> SEQUENCE: 1  
30 gggcaaaact cttccccccg ggatttatat gggaaagggg aaactttggc ccgtattcaa 60  
31 ggcgcacggg ttttggggcg gaatgaattt tttcgttccg gaaaaagtaa ttccccggga 120  
32 acgtagggta tcggtttcat aggctcgcca aatgggatat aggtggaaag gtaaaaaaaa 180  
33 ctgagccaag caaaggatag agaagctctg taatcatcgc aggttaaagg ggggatgta 240  
34 ttttagcctg caaatagtgt aattattgga tcctgtaaag agaaaaggac gaatgcgctg 300  
35 aagataagaa catttattga tattaataa ttaatttttt atgaagcggg gtaattaatt 360  
36 ttatctctca gcttttgtgt gatgcaaagc tctttccata agttctttct ttcaatgatt 420  
37 ctagcttatt cttgctgctc tttaaatggg gggggatatg cagcagaaat catgggtcct 480  
38 caaggaattt acgatgggga gacgttaact gtatcatttc cctatactgt tataggagat 540  
39 ccgagtggga ctactgtttt ttctgcagga gagttaacat taaaaaatct tgacaattct 600  
40 attgcagctt tgcctttaag ttgttttggg aacttattag ggagttttac tgttttaggg 660  
41 agaggacact cgttgacttt cgagaacata cggacttcta caaatggggc agctctaagt 720  
42 aatagcgtg ctgatggact gtttactatt gagggtttta aagaattatc cttttccaat 780  
43 tgcaattcat tacttgccgt actgcctgct gcaacgacta ataagggtag ccagactccg 840  
44 acgacaacat ctacaccgtc taatggtact atttattcta aaacagatct tttgttactc 900  
45 aataatgaga agttctcatt ctatagtaat ttagtctctg gagatggggg agctatagat 960  
46 gctaagagct taacggttca aggaattagc aagctttgtg tcttccaaga aaatactgct 1020  
47 caagctgatg ggggagcttg tcaagtagtc accagtttct ctgctatggc taacgaggct 1080  
48 cctattgcct ttgtagcgaa tggtgcagga gtaagagggg gagggattgc tgctgttcag 1140  
49 gatgggcagc agggagtgtc atcatctact tcaacagaag atccagtagt aagtttttcc 1200  
50 agaaatactg cggtagagtt tgatgggaac gtagcccgag taggaggagg gatttactcc 1260  
51 tacgggaacg ttgctttcct gaataatgga aaaaccttgt ttctcaacaa tgttgcttct 1320  
52 cctgtttaca ttgctgctaa gcaaccaaca agtggacagg cttctaatac gagtaataat 1380  
53 tacggagatg gaggagctat cttctgtaag aatggtgcgc aagcaggatc caataactct 1440  
54 ggatcagttt cctttgatgg agagggagta gttttcttta gtagcaatgt agctgctggg 1500  
55 aaagggggag ctatttatgc caaaaagctc tcggttgcta actgtggccc tgtacaattt 1560  
56 ttaaggaata tcgctaata tggtggagcg atttatttag gagaatctgg agagctcagt 1620  
57 ttatctgctg attatggaga tattattttc gatgggaatc ttaaaagaac agccaaagag 1680  
58 aatgctgccg atgttaatgg cgtaactgtg tcctcacaag ccatttcgat gggatcggga 1740

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/542,520

DATE: 05/31/2001

TIME: 11:51:53

Input Set : A:\7969.076.999.txt

Output Set: C:\CRF3\05312001\I542520.raw

```

59 gggaaaataa cgacattaag agctaaagca gggcatcaga ttctctttaa tgatcccatc 1800
60 gagatggcaa acggaaataa ccagccagcg cagtcttcca aacttctaaa aattaacgat 1860
61 ggtgaaggat acacagggga tattgttttt gctaattgaa gcagtacttt gtacccaaat 1920
62 gttacgatag agcaaggaag gattgttctt cgtgaaaagg caaaattatc agtgaattct 1980
63 ctaagtcaga caggtgggag tctgtatatg gaagctggga gtacatggga ttttgtaact 2040
64 ccacaaccac cacaacagcc tcctgccgct aatcagttga tcacgctttc caatctgcat 2100
65 ttgtctcttt cttctttgtt agcaaacaat gcagttacga atcctcctac caatcctcca 2160
66 gcgcaagatt ctcatcctgc agtcattggg agcacaactg ctggttctgt tacaattagt 2220
67 gggcctatct tttttgagga tttggatgat acagcttatg ataggtatga ttggctagggt 2280
68 tctaatacaa aaatcaatgt cctgaaatta cagttaggga ctaagccccc agctaattgcc 2340
69 ccatcagatt tgactctagg gaatgagatg cctaagtatg gctatcaagg aagctggaag 2400
70 cttgcgtggg atcctaatac agcaaataat ggtccttata ctctgaaagc tacatggact 2460
71 aaaactgggt ataatcctgg gcctgagcga gtagcttctt tggttccaaa tagtttatgg 2520
72 ggatccattt tagatatacg atctgcgcat tcagcaattc aagcaagtgt ggatgggcgc 2580
73 tcttattgtc gaggattatg ggtttctgga gtttcgaatt tcttctatca tgaccgcgat 2640
74 gctttagggtc agggatatcg gtatattagt gggggttatt ccttaggagc aaactcctac 2700
75 tttggatcat cgatgtttgg tctagcattt accgaagtat ttggtagatc taaagattat 2760
76 gtagtgtgtc gttccaatca tcatgcttgc ataggatccg tttatctatc taccacaaca 2820
77 gctttatgtg gatcctatct gttcggagat gcgtttatcc gtgctagcta cgggtttggg 2880
78 aatcagcata tgaaaacctc atatacattt gcagaggaga gcgatgttcg ttgggataat 2940
79 aactgtctgg ctggagagat tggagcggga ttaccgattg tgattactcc atctaagctc 3000
80 tatttgaatg agttgcgtcc tttcgtgcaa gctgagtttt cttatgccga tcatgaatct 3060
81 tttacagagg aaggcgatca agctcgggca ttcaagagcg gacatctcct aaatctatca 3120
82 gttcctgttg gagtgaagtt tgatcgatgt tctagtagac atcctaataa atatagcttt 3180
83 atggcggcct atatctgtga tgcttatcgc accatctctg gtactgagac aacgctccta 3240
84 tcccatcaag agacatggac aacagatgcc tttcatttag caagacatgg agttgtgggt 3300
85 agaggatcta tgtatgcttc tctaacaagt aatatagaag tatatggcca tgggaagatat 3360
86 gagtatcgag atgcttctcg aggctatggg ttgagtgcag gaagtagagt ccggttctaa 3420
87 aaatattggg tagatagtta agtgtagcgc atgccttttt ctttgagatc tacatcattt 3480
88 tgtttttttag cttgtttgtg ttcctattcg tatggattcg cgagctctcc tcaagtgtta 3540
89 acgcctaata taaccactcc ttttaaggga gacgatgttt acttgaatgg agactgcgct 3600
90 tttgtcaatg tctatgcagg agctgaagaa ggttcgatta tctcagctaa tggcgacaat 3660
91 ttaacgatta ccggacaaaa ccatacatta tcatttacag attctcaagg gccagttctt 3720
92 caaaattatg ctttcatttc agcaggagag acacttactc tgagagattt ttcgagtctg 3780
93 atgttctcga aaaatgtttc ttgcggagaa aagggaatga tctccgggaa aaccgtgagt 3840
94 atttccggag caggcgaagt gattttctgg gataactccg tggggtattc tcctttatct 3900
95 actgtgccaa cctcatcatc aactccgcct gctccaacag ttagtgatgc tcggaaaggg 3960
96 tctatttttt ctgtagagac tagtttgag atctcaggcg tcaaaaaagg ggtcatgttc 4020
97 gataataatg ccggaattt cggaacagtt tttcgaggta agaataataa taatgtcggg 4080
98 ggtggaggca gtgggttccg ctacaccatc aagtacgact ttacagtta aaaactgtaa 4140
99 agggaaagtt tctttcacag ataacgtagc ctcttcgga ggcggagtgg tttataaagg 4200
100 cattgtgctt ttcaaagaca atgaaggagg catattcttc cgagggaaca cagcatatca 4260
101 tgattttaagg attcttgctg ctactaatca ggatcagaat acggagacag gaggcggtgg 4320
102 aggagttatt tgctctccag atgattctgt aaagtttgaa ggcaataaag gttctattgt 4380
103 ttttgattac aactttgcaa aaggcagagg cggaagcatc ctaacgaaag aattc 4435
105 <210> SEQ ID NO: 2
106 <211> LENGTH: 1012
107 <212> TYPE: PRT
108 <213> ORGANISM: Chlamydia

```

## RAW SEQUENCE LISTING

DATE: 05/31/2001

PATENT APPLICATION: US/09/542,520

TIME: 11:51:53

Input Set : A:\7969.076.999.txt

Output Set: C:\CRF3\05312001\I542520.raw

110 &lt;400&gt; SEQUENCE: 2

```

111 Met Gln Thr Ser Phe His Lys Phe Phe Leu Ser Met Ile Leu Ala Tyr
112   1      5      10      15
114 Ser Cys Cys Ser Leu Asn Gly Gly Gly Tyr Ala Ala Glu Ile Met Val
115      20      25      30
117 Pro Gln Gly Ile Tyr Asp Gly Glu Thr Leu Thr Val Ser Phe Pro Tyr
118      35      40      45
120 Thr Val Ile Gly Asp Pro Ser Gly Thr Thr Val Phe Ser Ala Gly Glu
121      50      55      60
123 Leu Thr Leu Lys Asn Leu Asp Asn Ser Ile Ala Ala Leu Pro Leu Ser
124      65      70      75      80
126 Cys Phe Gly Asn Leu Leu Gly Ser Phe Thr Val Leu Gly Arg Gly His
127      85      90      95
129 Ser Leu Thr Phe Glu Asn Ile Arg Thr Ser Thr Asn Gly Ala Ala Leu
130      100      105      110
132 Ser Asn Ser Ala Ala Asp Gly Leu Phe Thr Ile Glu Gly Phe Lys Glu
133      115      120      125
135 Leu Ser Phe Ser Asn Cys Asn Ser Leu Leu Ala Val Leu Pro Ala Ala
136      130      135      140
138 Thr Thr Asn Lys Gly Ser Gln Thr Pro Thr Thr Ser Thr Pro Ser
139 145      150      155      160
141 Asn Gly Thr Ile Tyr Ser Lys Thr Asp Leu Leu Leu Leu Asn Asn Glu
142      165      170      175
144 Lys Phe Ser Phe Tyr Ser Asn Leu Val Ser Gly Asp Gly Gly Ala Ile
145      180      185      190
147 Asp Ala Lys Ser Leu Thr Val Gln Gly Ile Ser Lys Leu Cys Val Phe
148      195      200      205
150 Gln Glu Asn Thr Ala Gln Ala Asp Gly Gly Ala Cys Gln Val Val Thr
151      210      215      220
153 Ser Phe Ser Ala Met Ala Asn Glu Ala Pro Ile Ala Phe Val Ala Asn
154 225      230      235      240
156 Val Ala Gly Val Arg Gly Gly Gly Ile Ala Ala Val Gln Asp Gly Gln
157      245      250      255
159 Gln Gly Val Ser Ser Ser Thr Ser Thr Glu Asp Pro Val Val Ser Phe
160      260      265      270
162 Ser Arg Asn Thr Ala Val Glu Phe Asp Gly Asn Val Ala Arg Val Gly
163      275      280      285
165 Gly Gly Ile Tyr Ser Tyr Gly Asn Val Ala Phe Leu Asn Asn Gly Lys
166      290      295      300
168 Thr Leu Phe Leu Asn Asn Val Ala Ser Pro Val Tyr Ile Ala Ala Lys
169 305      310      315      320
171 Gln Pro Thr Ser Gly Gln Ala Ser Asn Thr Ser Asn Asn Tyr Gly Asp
172      325      330      335
174 Gly Gly Ala Ile Phe Cys Lys Asn Gly Ala Gln Ala Gly Ser Asn Asn
175      340      345      350
177 Ser Gly Ser Val Ser Phe Asp Gly Glu Gly Val Val Phe Phe Ser Ser
178      355      360      365
180 Asn Val Ala Ala Gly Lys Gly Gly Ala Ile Tyr Ala Lys Lys Leu Ser
181      370      375      380

```

## RAW SEQUENCE LISTING

DATE: 05/31/2001

PATENT APPLICATION: US/09/542,520

TIME: 11:51:53

Input Set : A:\7969.076.999.txt

Output Set: C:\CRF3\05312001\I542520.raw

```

183 Val Ala Asn Cys Gly Pro Val Gln Phe Leu Arg Asn Ile Ala Asn Asp
184 385 390 395 400
186 Gly Gly Ala Ile Tyr Leu Gly Glu Ser Gly Glu Leu Ser Leu Ser Ala
187 405 410 415
189 Asp Tyr Gly Asp Ile Ile Phe Asp Gly Asn Leu Lys Arg Thr Ala Lys
190 420 425 430
192 Glu Asn Ala Ala Asp Val Asn Gly Val Thr Val Ser Ser Gln Ala Ile
193 435 440 445
195 Ser Met Gly Ser Gly Gly Lys Ile Thr Thr Leu Arg Ala Lys Ala Gly
196 450 455 460
198 His Gln Ile Leu Phe Asn Asp Pro Ile Glu Met Ala Asn Gly Asn Asn
199 465 470 475 480
201 Gln Pro Ala Gln Ser Ser Lys Leu Leu Lys Ile Asn Asp Gly Glu Gly
202 485 490 495
204 Tyr Thr Gly Asp Ile Val Phe Ala Asn Gly Ser Ser Thr Leu Tyr Gln
205 500 505 510
207 Asn Val Thr Ile Glu Gln Gly Arg Ile Val Leu Arg Glu Lys Ala Lys
208 515 520 525
210 Leu Ser Val Asn Ser Leu Ser Gln Thr Gly Gly Ser Leu Tyr Met Glu
211 530 535 540
213 Ala Gly Ser Thr Trp Asp Phe Val Thr Pro Gln Pro Pro Gln Gln Pro
214 545 550 555 560
216 Pro Ala Ala Asn Gln Leu Ile Thr Leu Ser Asn Leu His Leu Ser Leu
217 565 570 575
219 Ser Ser Leu Leu Ala Asn Asn Ala Val Thr Asn Pro Pro Thr Asn Pro
220 580 585 590
222 Pro Ala Gln Asp Ser His Pro Ala Val Ile Gly Ser Thr Thr Ala Gly
223 595 600 605
225 Ser Val Thr Ile Ser Gly Pro Ile Phe Phe Glu Asp Leu Asp Asp Thr
226 610 615 620
228 Ala Tyr Asp Arg Tyr Asp Trp Leu Gly Ser Asn Gln Lys Ile Asn Val
229 625 630 635 640
231 Leu Lys Leu Gln Leu Gly Thr Lys Pro Pro Ala Asn Ala Pro Ser Asp
232 645 650 655
234 Leu Thr Leu Gly Asn Glu Met Pro Lys Tyr Gly Tyr Gln Gly Ser Trp
235 660 665 670
237 Lys Leu Ala Trp Asp Pro Asn Thr Ala Asn Asn Gly Pro Tyr Thr Leu
238 675 680 685
240 Lys Ala Thr Trp Thr Lys Thr Gly Tyr Asn Pro Gly Pro Glu Arg Val
241 690 695 700
243 Ala Ser Leu Val Pro Asn Ser Leu Trp Gly Ser Ile Leu Asp Ile Arg
244 705 710 715 720
246 Ser Ala His Ser Ala Ile Gln Ala Ser Val Asp Gly Arg Ser Tyr Cys
247 725 730 735
249 Arg Gly Leu Trp Val Ser Gly Val Ser Asn Phe Phe Tyr His Asp Arg
250 740 745 750
252 Asp Ala Leu Gly Gln Gly Tyr Arg Tyr Ile Ser Gly Gly Tyr Ser Leu
253 755 760 765
255 Gly Ala Asn Ser Tyr Phe Gly Ser Ser Met Phe Gly Leu Ala Phe Thr

```

## RAW SEQUENCE LISTING

DATE: 05/31/2001

PATENT APPLICATION: US/09/542,520

TIME: 11:51:53

Input Set : A:\7969.076.999.txt

Output Set: C:\CRF3\05312001\I542520.raw

```

256      770      775      780
258 Glu Val Phe Gly Arg Ser Lys Asp Tyr Val Val Cys Arg Ser Asn His
259 785      790      795      800
261 His Ala Cys Ile Gly Ser Val Tyr Leu Ser Thr Gln Gln Ala Leu Cys
262      805      810      815
264 Gly Ser Tyr Leu Phe Gly Asp Ala Phe Ile Arg Ala Ser Tyr Gly Phe
265      820      825      830
267 Gly Asn Gln His Met Lys Thr Ser Tyr Thr Phe Ala Glu Glu Ser Asp
268      835      840      845
270 Val Arg Trp Asp Asn Asn Cys Leu Ala Gly Glu Ile Gly Ala Gly Leu
271      850      855      860
273 Pro Ile Val Ile Thr Pro Ser Lys Leu Tyr Leu Asn Glu Leu Arg Pro
274 865      870      875      880
276 Phe Val Gln Ala Glu Phe Ser Tyr Ala Asp His Glu Ser Phe Thr Glu
277      885      890      895
279 Glu Gly Asp Gln Ala Arg Ala Phe Lys Ser Gly His Leu Leu Asn Leu
280      900      905      910
282 Ser Val Pro Val Gly Val Lys Phe Asp Arg Cys Ser Ser Thr His Pro
283      915      920      925
285 Asn Lys Tyr Ser Phe Met Ala Ala Tyr Ile Cys Asp Ala Tyr Arg Thr
286      930      935      940
288 Ile Ser Gly Thr Glu Thr Thr Leu Leu Ser His Gln Glu Thr Trp Thr
289 945      950      955      960
291 Thr Asp Ala Phe His Leu Ala Arg His Gly Val Val Val Arg Gly Ser
292      965      970      975
294 Met Tyr Ala Ser Leu Thr Ser Asn Ile Glu Val Tyr Gly His Gly Arg
295      980      985      990
297 Tyr Glu Tyr Arg Asp Ala Ser Arg Gly Tyr Gly Leu Ser Ala Gly Ser
298      995      1000      1005
300 Arg Val Arg Phe
301      1010

```

304 &lt;210&gt; SEQ ID NO: 3

305 &lt;211&gt; LENGTH: 20

306 &lt;212&gt; TYPE: PRT

307 &lt;213&gt; ORGANISM: Chlamydia

309 &lt;220&gt; FEATURE:

310 &lt;221&gt; NAME/KEY: SITE

311 &lt;222&gt; LOCATION: 19

312 &lt;223&gt; OTHER INFORMATION: Xaa=unknown amino acid

314 &lt;400&gt; SEQUENCE: 3

315 Glu Ile Met Val Pro Gln Gly Ile Tyr Asp Gly Glu Thr Leu Thr Val

316 1 5 10 15

318 Ser Phe Xaa Tyr

319 20

322 &lt;210&gt; SEQ ID NO: 4

323 &lt;211&gt; LENGTH: 18

324 &lt;212&gt; TYPE: DNA

325 &lt;213&gt; ORGANISM: Chlamydia

327 &lt;220&gt; FEATURE:

**Please Note:**

Use f n and/or Xaa hav been detected in the Sequ nce Listing. Please review the Sequence Listing to ensure that a corresp nding explanation is presented in the <220> to <223> fields of each sequence which pres nts at least one n or Xaa.

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/542,520

DATE: 05/31/2001

TIME: 11:51:54

Input Set : A:\7969.076.999.txt

Output Set: C:\CRF3\05312001\I542520.raw

L:318 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:334 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:347 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:361 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6  
L:374 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7  
L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8  
L:400 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9